

## THE STRUCTURE OF THE KHWAREZMIAN VERB

by W. B. HENNING

Our<sup>1</sup> knowledge of the Khwarezmian language depends largely on two classes of material, which are dissimilar in character but supplement each other. On the one hand, we have the sentences in legal books, chief among them the *Qunyatu 'lMunyah*; on the other, the Khwarezmian glosses in the *Muqaddimatu 'lAdab*, the principal MS. of which was published in facsimile by Zeki Velidi Togan three years ago.<sup>2</sup>

In the *Muqaddimah* the translator aims at accurately reproducing the Arabic original. His phrases are unidiomatic, cast in a uniform mould, monotonous, repetitive, and deadly dull; but for us they are highly informative, and would be even more so if the scribe had not chosen to omit for long stretches the diacritical points, without which Arabic script is difficult to read even if one already knows the language. To give simple examples, he does not mind writing سك for "house", which admits of 36 readings,<sup>3</sup> or سد for "he became", which in theory can be read in 2640 different ways.<sup>4</sup> However, words which are left unpointed in one place are often found pointed in another, so by careful comparison the correct reading can be established for the greater part of the material. I have now compiled nearly a complete glossary which I hope to publish in the near future.

The Khwarezmian sentences in the *Qunyatu 'lMunyah*, on the contrary, are highly idiomatic, often slangy, full of puns and double meanings, which are difficult to grasp unless one is already thoroughly acquainted with the language. These sentences come from case law; they are sentences actually used in life, which subsequently acquired significance in a law suit. If then we wish to study the syntax of the Khwarezmian language and its mode of expression in daily life, we shall have to turn to the *Qunyah*; while the *Muqaddimah* serves largely to complete our dictionary. Moreover, the MSS. of the *Qunyah* have all the diacritical marks one could wish for, in fact

<sup>1</sup> Paper read to the Iranian Section of the xxiii<sup>rd</sup> International Congress of Orientalists, on 25 August 1954.

<sup>2</sup> *Khorezmian Glossary of the Muqaddimat al-Adab*, Istanbul, 1951.

<sup>3</sup> Recte *pðk* (پځ), perhaps from *paða-* with change from -ð- to -ð- (as in Younger Avestan).

<sup>4</sup> Recte *nyðyd* (نځيد), 3rd sing. imperfect (pause form) of *nyð-* "to sit, to become" (Sogd. *nyð-*).

more than one would like to have; there is a profusion of diacritical points, and vowel marks in addition, but unfortunately each MS. is apt to have somewhat different ones from the next.

Eighteen years ago I gave a brief report on the Khwarezmian language, based largely on the *Qunyah*.<sup>1</sup> At the time I had seen the material only for a few weeks, and so my report embodied some misapprehensions and omitted some important points. Subsequent publications from another side, based also on the *Qunyah*, have neither removed the misapprehensions nor filled in any of the gaps.<sup>2</sup> To-day I would say that precisely the most interesting and important points remained unrecognized, and that is true also of the structure of the Khwarezmian verb; some of them I should like to bring to your attention.

In the *Muqaddimah* the dominant verbal form is the 3rd person singular of the imperfect; next in frequency is the 3rd plural, also of the imperfect; all other forms are very rare. This restriction, although obviously a disadvantage in some ways, is in fact beneficial on the whole; for it allows us to observe, at one example, the infinitely complicated construction of verbal forms with clarity and thus enables us to understand also the rarer forms for the other persons, tenses and moods.

The 3rd singular of the imperfect appears with four endings, *-d*, *-yd*, *-yyd*, and *-yt*. They prove the existence of three classes of present stems.

The first endings, *-d* and *-yd*, are variants of one and the same form. Any verb the stem of which ends in a consonant possesses these two variants. For example, from *hβr-* (*hiβr-*) "to give" we have *h'βrd* and *h'βryd* "he gave". The distribution of these two forms depends on the position in the sentence: the longer form is confined to the end of a sentence, or, to borrow a term from Hebrew grammar, *stands in pause*.

The existence of pause forms is one of the most interesting facts about the Khwarezmian language. Generally, the vowel before the last consonant of a word is stressed in pause and thereby lengthened or apparently lengthened. This rule applies to all words; accordingly, a noun such as *zādik* "son" appears in two forms, as *z'dk* and, in pause, as *z'dyk*; this represents Old Iranian *zātaka-*, for Old Iranian interior *-a-* generally becomes *-i-* in Khwarezmian.

Our example, the word for "he gave", can then be accounted for in this way: it represents *frābarata*; the shorter form, (*h'βrd*), pronounced *hāβirda*, is *frābarata*, but the pause form (*h'βryd*), apparently pronounced *hāβrīda*,

<sup>1</sup> ZDMG., 90 (1936), pp. \*30\*—\*34\*.

<sup>2</sup> An evaluation of the various contributions to Khwarezmian studies was given in an article, written in 1950, which will be included in the *Mélanges Z.V. Togan* (Z.V. Togan Armağanı), pp. 421–36, the publication of which has regrettably been delayed. That article also contains a sketch of the phonology, a discussion of grammatical points, and a list of interesting words. For the time being see my *Zoroaster—politician or witch-doctor?*, 44 sq.

is *frābarata*. This distinction must have existed already in the Old Iranian form of Khwarezmian.

Sometimes the verbal stems are affected. This happens chiefly in the imperative. For example, the most frequent word in our material is *p'rwzd* (*pārwuzda*) "it became". Its present stem is *prwz-*, probably reflecting Old Iranian *pari-waza-*; but the imperative, in pause, is *prwyz* (*pirwīza*). Similarly, "to eat" is *xr-*, vocalized *xur-*; but the imperative is *xwyr* = *xwīra*.<sup>1</sup>

To return to the endings of the 3rd singular of the imperfect, *-yd* is also the interior form of stems ending in *-y*; in such cases, *-yd* has beside it a pause form in *-yyd*. Stems in *-y* are very common; they mostly continue ancient causative stems in *-aya*; but sometimes they represent presents from bases in *-ī*. For example, *mxyd*, in pause *mxyyd*, "he wept or mourned", pl. *mxy'r*, which derives from *xšī*; the *m*-prefix proves that the stem began with two consonants.<sup>2</sup>

In the fourth ending, *-yt*, *-y-* again forms part of the stem. The *-t* was preserved, instead of changing into *-d*, either because no vowel preceded it, or because a preceding vowel was lost at an early stage. These imperfects belong to bases in *-ā*, in *-ah*, and to the base *i* "to go" if joined with a preverb; the simple *i* has *myd* (*mīda*), in pause *myyd* (*miyīda*), "he went". For example, *fra-mā* "to command" produces *š'myt* (*šāmayta*) and a present stem *šm'h-*,<sup>3</sup> with secondary *-h*; from *ah-* "to be" we have *myt* (*mayta*) "he was". "To enter" is *cy-* (*cīyy-*) in the present, which I once wrongly explained as from *ēyawa-*, while in fact it is *ati* + *i*; its imperfect is *c'yt* (*cāyta*). These verbs are further peculiar in having *-la* as ending of the 3rd plural, e.g., *c'yl* (*cāyla*) "they entered", while all other verbs have *-āra* in the 3rd plural of the imperfect. The same opposition *-t* : *-l-* is found also in the present indicative of "to be", *yt* (*yetti*) "he is", *yl* (*yelli?*) "they are". The whole group doubtless represents the remains of the ancient non-thematic inflexion.

<sup>1</sup> This, therefore, represents *xwara*, not by any means \**xwaraya* (as one might imagine, perhaps misled by Sogdian).

<sup>2</sup> Only verbs that begin with a vowel in the present stem have the *m*-prefix in the imperfect. The present stem to which *mxy(y)d* belonged was therefore \**xy-*. The initial (prosthetic) vowel arose because the stem originally began with two consonants (*xf*), which were later simplified (*xx*).

<sup>3</sup> In the discussion following the reading of this paper I was invited to account for the divergent treatment of *fra-* exemplified by *šm'h-* (if from *fra-mā-*) and *hβr-* (if from *fra-bar-*). In fact, the preposition *fra-* in verbal compounds normally became *ša-* in Khwarezmian (there are many examples); *hβr-* is as much an exception in Khwarezmian as the corresponding verb, *ββr-*, is in Sogdian; for an explanation see I. Gershevitch, *JRAS.*, 1946, 181, n. 3; *Gramm. Manich. Sogdian*, p. 246. On the history of *fr* in Khwarezmian in general see *Mélanges Z.V. Togan*, pp. 433 sq. One of the words quoted there in illustration, *šw'h* "pure" as from \**fra-pawāka-* (with *fra-* conferring elative meaning, comparable to \**peri-* in other languages, see Ernout-Meillet<sup>2</sup>, 880 sq.), is somewhat uncertain as regards its origin; for its *š-* might respond to Avestan *šraf-*.

Having discussed the three present classes revealed by the four endings, I want now to draw your attention to their finals. They seem to end in a consonant, *d* or *t*, but in reality end in a vowel, in *da* or *ta*. The existence of such unwritten final vowels is assured, not so much by the somewhat erratic vocalization in the *Qunyah*, as by the behaviour of the endings when a suffix is added. Then the final vowel is apparently lengthened, or at any rate expressed in writing as if it were a long vowel.

Most verbal forms, and most nominal forms too,<sup>1</sup> possess final brief vowels, generally *-a* or *-i*, which appear in writing only before suffixes. The question to what extent these final vowels continue the corresponding Old Iranian endings is of great complexity and cannot be discussed here; in the case of the 3rd singular of the imperfect there is no difficulty: it agrees perfectly with the Old Iranian middle ending, *-ta*.

The most common suffixes are the enclitic pronouns for the 3rd person, *-hi* in the singular, *-na* and *-hina* in the plural, the former for the accusative, the latter for the general oblique case. That *-na*, which has a close relative in Khotanese,<sup>2</sup> is the accusative<sup>3</sup> of the plural only is not immediately obvious; it often refers to nouns seemingly in the singular, which, however, are plurals in fact, collectives, the terms for water, food, grain, wheat, and the like.<sup>4</sup>

These suffixes, and most others as well, end in their turn in brief vowels which are not expressed in the orthography unless a further suffix is added. For example, the normal equivalent of "he gave" is *k'brd* = *hābirda*; "he gave him" is *hābirdāhi*, spelt *k'brd'h*; "he gave them to him" is *hābirdāhina*, spelt *k'brd'hyn*<sup>5</sup>; if a further suffix is added, the final *-a* will be expressed in writing, therefore *k'brd'hyn'*.

It is a rule that any vocalic ending preserves its vowel unchanged, irrespective of the sounds of the suffix. There is only one exception to this

<sup>1</sup> With few exceptions, chief among them forms of masculine nouns in the singular. Although it is too early yet to be sure of all points, the following scheme of inflection in the singular may be put forward tentatively:

	Masc.	Fem.	Fem. <i>-k</i> stems
Nom.-Acc.			
Genitive (a)	} nil	<i>-a</i>	<i>-ka</i>
Possessive			
Ablative (b)	} <i>-ān</i>	} <i>-ya</i>	} <i>-ca</i>
Locative (c)			
	} <i>-a</i>	<i>-a</i>	<i>-ka</i>

(a) Chiefly before the postposition *δ'r*. (b) After *c*- "from". (c) Chiefly after the prepositions *f*- "in" and *par*- "on".

<sup>2</sup> In Sogdian, differently, *-n* serves for both numbers, cf. *BSOAS.*, xii, 605, n. 3.

<sup>3</sup> The distinction between *-na* and *-hina* is strict, though obscured by the occasional loss of *-h* in *-hina* (with attendant elision of an antecedent vowel) and by the possible confusion of *-hina* with *-hi-na* (the two pronouns, *-hi* and *-na*, in succession).

<sup>4</sup> In this point Khwarezmian strikingly resembles Pashto (cf. Trumpp §53).

<sup>5</sup> This, of course, could equally well mean "he gave to them".

rule: namely, when the suffixes *-hi* or *-hina* are followed by a further suffix, they may lose their *-h*-, whereupon the preceding vowel disappears. Thus, instead of *hābirdāhina* one may say, and usually does, *hābirdina*, spelt *k'brdyn*.

We must now consider the suffixes that may be attached to verbal forms. They fall into three classes. *Firstly*, the enclitic pronouns, nine in number, three of them have already been mentioned, the remaining six are for the 1st and 2nd persons. *Secondly*, certain postpositions, principally the following four: *c'* "from"<sup>1</sup> Old Iranian *hačā*; *δa* "with, to" Old Iranian *hada*; *b'r* "upon" Old Iranian *upari*; and *da*, with ill-defined meaning (approximately "through, beyond, off"), probably connected with Old Iranian *ati*. *Thirdly*, two adverbs with local meaning, *wa* and *w'θ*, the latter from Old Iranian *awada*.

When several suffixes are attached to a verb, they observe a fixed order of precedence. The pronouns come first, then postpositions and adverbs. Within each of these two groups, the sequence is again strict. In the pronouns, for example, *hi* precedes *na* and *di*, but follows *mi*; among the postpositions and adverbs, *wa* follows *ci*, but precedes *bir*; and so on.

The sequence is often in conflict with the meaning. The postpositions may modify the verbal meaning; or they may function as directives to preceding enclitic pronouns. In the latter case it happens that the enclitic pronoun is separated from the postposition by some other suffix, often one expressing the direct object of the verb. That produces weird forms, e.g., *hyδdyn*<sup>2</sup> "he read them to him", *c'yt'hyw'byr* "he entered there before him".<sup>3</sup> This conflict of meaning and fixed sequence can sometimes be observed also when enclitics are added to words other than verbal forms, for example: *kf'n'c m'sn* "that I took them from you".<sup>4</sup>

As we have seen, the suffixes that may be added to verbs number at least fifteen. If one takes into account that a given verbal form is capable of taking as many as four different suffixes at a time, it will become clear that there is an almost endless variety of possible forms. Some are veritable

<sup>1</sup> *-cī* (*-cy-*) when followed by another suffix; but *cā* (*c'*) in independent position, as adverb (e.g., *cā wāzayta* or *wāzaytdci* "he went out"). Not to be confused with the preposition *c*- which, as a rule, appears in combination with the article (*cī* "from the" masc. and pl. comm., *cā* fem.)

<sup>2</sup> = *hyδd'-hy-n'-θ* "he read-him-them (acc.)-to". The origin of *hyδ-* "to read, recite" is uncertain. Formally, *hyδ-* may easily represent OIr. *hida-* (Skt. *sida-*) "to sit"; a special development of the meaning (*assideo*, to sit in on—perhaps—a sacrifice and recite hymns?) seems not impossible. Connection with Av. *ašaya-* is unlikely.

<sup>3</sup> = *c'yt'-hy-w'-byr* "he entered-him-there-upon".

<sup>4</sup> "That-you-them-from I took" ("from you" is *f'c*). Thanks to its position at the end, the verb is devoid of suffixes; otherwise, one would doubtless say *\*m'sn'f'n'c* (the corresponding 3rd person is *\*m'sd'f'n'c*, or rather, with assimilation, *\*m'st'f'n'c* "he took them from you").

monsters, such as *hyδd'hyn'd'br* "he recited them before him".<sup>1</sup> No other Iranian language knows such a heaping of suffixes; whether Khwarezmian has been influenced by Turkish in this respect, it is impossible to tell so long as the pre-Muslimic documents remain undeciphered.

Now I come to my last point, the *principle of anticipation*. The position of the verb in a sentence is fairly free. If it precedes the object, the object is expressed a second time by an enclitic pronoun either on the verb itself or on a word before the verb. Until one grasps this principle, it is impossible to understand the structure of the Khwarezmian verbal forms. A good example is provided by the first Khwarezmian sentence I quoted eighteen years ago,<sup>2</sup> at a time when this point had not been clear to me. It meant "I have given you my daughter for one hundred pieces of gold" and began with the verb: *h'βrnyd y' δwyd'm* = *hāβirnīdi yā δuydāmi* "I have given you my daughter". At the time I thought that *hāβirnīdi* was the 1st person of the imperfect plus the enclitic pronoun for the 2nd person, genitive-dative, that is *di*. That was incorrect. Firstly, the 1st person of the imperfect ends in *-(i)na*, so that "I gave" is *hāβirna*, or in pause *hāβirna*; and if *-di* is added, the result is necessarily *hāβirnādi*, not *hāβirnīdi*. Secondly, in a sentence of this type the object *must* be expressed on the verb by the relevant enclitic pronoun, in this case *-hi*. Thus, *hāβirnīdi* is necessarily equal to *\*hāβirnā-hīdi*, and the sentence means literally "I have given *her* to you, my daughter".

Most of the longer verbal forms contain one or even two anticipatory suffixes; it is in fact largely due to them that the forms have become so inconveniently long. Not only the direct object, but also other parts of the sentence, such as prepositional phrases, may be expressed on a preceding verb by anticipatory suffixes. For example, *w'xyt'c c'γdc 'y xr'x* (*wāzaytāci cāyudca ī xērāx*) "the sword was drawn from the scabbard, lit. came out".<sup>3</sup> Here the postposition *ci*, attached to the verb, anticipates the immediately following preposition; if the order of the words is changed and the verb placed at the end, it will appear as *w'xyt*. Similarly, the adverb *wa* "there(in)" often anticipates the preposition *fi* "in". Thus, *m'st'n'w fy prδ'β* "he misled them",<sup>4</sup> lit. "he took them (there)in, in deception"; or *m'kwcydyw fycwb* "he dipped it in water".<sup>5</sup>

<sup>1</sup> "He read-him-them-off-upon" (*-d'* serves to modify the verbal meaning, *hyδ-* + *d'* approximately = "recite").

<sup>2</sup> ZDMG. 90 (1936), p. \*34\*.

<sup>3</sup> *wāzayta* "it came out" (from *uz-i*, treated in the imperfect as if it were compounded with *awa-*); *'γdyk* fem. "scabbard", gen.-abl. *'γdc*; *xr'x* "sword" cf. Sogd. *xyr*, etc.

<sup>4</sup> *māstā-nā-wa* from *ās-* "to take", impf. 3rd sing. in pause *m'ryd* = *māsida*, otherwise *m'sd*, *m'st* = *māsda*, *māsta* (cf. above p. 47, n. 4).

<sup>5</sup> Short for *m-ākūcidā-hi-wa*; *ākūciy-*, caus. "to suspend, submerge".

These, then, are the four points that dominate, so it seems to me, the structure of the Khwarezmian verb: the presence of pause forms, the vocalic endings, the system of suffixes and its complications, and finally the existence of anticipatory suffixes.<sup>1</sup>

<sup>1</sup> It should perhaps be stated that the contents of this paper are largely inferential, based on imperfect and often contradictory material; in particular, the vocalization of the Khwarezmian words is almost wholly conjectural.